

Geometría Computacional y Bases de Datos

Susana Esquivel, Edilma Olinda Gagliardi, Maria Gisela Dorzán, Maria Teresa Taranilla,

Pablo Rafael Palmero y Carlos Andrés Casanova

Departamento de Informática
Facultad de Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis, Argentina
{esquivel, oli, mgdorzan, tarani, prpalmero}@unsl.edu.ar

Gregorio Hernández Peñalver

Departamento de Matemática Aplicada
Facultad de Informática
Universidad Politécnica de Madrid, España
gregorio@fi.upm.es

Resumen

Consideramos la línea de investigación denominada *Geometría Computacional y Bases de Datos* del proyecto *Tecnologías Avanzadas de Bases de Datos*, orientada a vincular las disciplinas Bases de Datos, Geometría Computacional y Metaheurísticas. El objetivo general consiste en utilizar métodos y herramientas de estas disciplinas para investigación de base en la resolución de problemas NP orientados a optimización, y también para la resolución de problemas en diversos dominios de aplicación.

Palabras clave: Bases de Datos, Geometría Computacional, Metaheurísticas, Bases de Datos Espaciales y Espacio-Temporales.

Contexto

El Proyecto Tecnologías Avanzadas de Bases de Datos desarrolla actividades vinculadas al tratamiento de objetos de diversos tipos, estructurados y no estructurados que son de utilidad en diversos campos de aplicación, tales como sistemas de información geográfica, computación gráfica, computación móvil, robótica, diseño asistido por computadora, motores de búsqueda en internet, entre otras, y que se relacionan en tales bases de datos.

En el proyecto existen tres líneas de investigación, orientadas al desarrollo de nuevos modelos para administrar y recuperar información almacenada en repositorios de datos no estructurados, donde los escenarios de exploración requieren modelos tales como las bases de datos de texto, bases de datos espaciales, espacio-temporales, bases de datos de imágenes, bases de datos de sonidos, espacios métricos, entre otros.

Así, surge el estudio de modelos como las bases de datos espaciales y bases de datos espacio-temporales. También, la necesidad de construir y manipular diferentes objetos y estructuras geométricas útiles en diversas áreas de aplicación.

En particular, las estructuras geométricas que se estudian deben cumplir con propiedades deseables, y algunos de los problemas relacionados con la optimización de las mismas son problemas NP-duros, por tanto en la búsqueda de soluciones aproximadas se aplican metaheurísticas.

Por otro lado, con el avance de las tecnologías, situaciones del mundo real ameritan el uso de las mismas, en pos de una mejora sustancial en cuestiones de calidad, gestión, economía, etc. A nivel mundial, las tecnologías de la información y la comunicación (TIC) continúan introduciendo cambios políticos,

económicos, sociales y culturales, entre muchos otros, convirtiéndose en procesos clave para cualquiera que quiera accionar sobre el presente y proyectar hacia el futuro.

Por lo expuesto, en la línea de investigación *Geometría Computacional y Bases de Datos* se vinculan temáticas que surgen de las disciplinas Bases de Datos, Geometría Computacional y Metaheurísticas, a fin de tratar con la optimización en problemas NP, o bien de poder trabajar en dominios de aplicación con soluciones integradas que hagan uso de técnicas y herramientas de estas disciplinas.

El trabajo de investigación se desarrolla en forma conjunta con investigadores afines de proyectos de esta Universidad, o de alianzas y convenios entre organizaciones nacionales y provinciales con presencia en San Luis, como así también de universidades extranjeras (Universidad Politécnica de Madrid - España, Universidad Veracruzana - México, Universidad del Bío Bío- Chile, entre otras) mediante convenios de cooperación interinstitucional.

En este trabajo, se exponen los tópicos en estudio, junto con las propuestas más recientes de interés.

1. Introducción

La optimización es una línea de investigación en Ciencias de la Computación, donde se procura encontrar la mejor solución posible a un problema dentro de un período de tiempo limitado. En el caso particular de problemas de optimización combinatoria, los hay NP-duros y los hay polinómicos; y no se puede garantizar encontrar la mejor solución en un tiempo razonable, para todas las instancias del problema.

En Geometría Computacional, la optimización de configuraciones geométricas respecto de ciertos criterios de calidad, pertenecen a esta clase de problemas, y pueden resolverse utilizando métodos de aproximación, tales como las técnicas metaheurísticas [BCKO08].

Una metaheurística es un proceso de generación iterativo que guía la búsqueda de soluciones combinando inteligentemente diferentes conceptos de campos diversos como inteligencia artificial, evolución biológica, inteligencia colectiva, sistemas inmunes, entre otros [MF04].

Entre los objetivos de la línea está proponer soluciones aproximadas para problemas geométricos para los cuales aún no se han encontrado algoritmos eficientes que los solucionen debido a su complejidad, mediante la aplicación de técnicas metaheurísticas.

Las propuestas de estudio están orientadas a la optimización de diferentes configuraciones geométricas, tales como triangulaciones y pseudotriangulaciones, poligonizaciones, entre otras. Los criterios de calidad considerados son peso, dilación, número de apuñalamiento mínimo, número de guardias en problemas de vigilancia, área, perímetro, entre otros. Estos criterios inducen a buscar soluciones óptimas respecto de ellos, mediante la aplicación de técnicas metaheurísticas.

Algunos de los problemas de optimización estudiados son la Triangulación de Peso Mínimo (*Minimum Weight Triangulation, MWT*) y la Pseudo-Triangulación de Peso Mínimo (*Minimum Weight Pseudo-Triangulation, MWPT*), problemas de carácter NP-duro [MR06], [GL07]. La Triangulación de Dilación Mínima (*Minimum Dilation Triangulation, MDT*) es otro problema estudiado, donde la dilación mide la calidad de conexión entre puntos de la triangulación. Con respecto a poligonizaciones, se estudió el problema de obtener poligonizaciones con mínima área para un conjunto de puntos en el plano [TGH11].

Otro tópico de interés actualmente en análisis, refiere a problemas de vigilancia. Los problemas de vigilancia pueden interpretarse, muchas veces, como problemas de iluminación o vigilancia. Se han planteado numerosas variantes del problema, cuestionándose ¿qué se vigila? y ¿cómo se vigila? De esta forma, se trata con diversos objetos geométricos a

vigilar y con diversas formas de vigilancia. En particular, nos interesan como objeto geométrico de estudio las triangulaciones planas las cuales constituyen un entorno geométrico adecuado, en el que también tienen sentido las preguntas sobre vigilancia. Un guardia situado en un vértice (ó arista) de la triangulación vigila todos los triángulos incidentes al vértice (ó arista). Por tanto, dada una triangulación T , podemos considerar las siguientes preguntas ¿cuántos guardias, ubicados en vértices (ó aristas), se necesitan para vigilar todos los triángulos de T ?, y ¿dónde se deben ubicar? es decir, cuáles son los vértices (ó aristas) seleccionados para ubicar los guardias.

Además de abordar el estudio de estos problemas mencionados, se busca poder hacer uso de soluciones en dominios de aplicación, naturales para su desarrollo.

La utilización de estas configuraciones geométricas optimizadas respecto de algún criterio de calidad, resultan como soporte de estrategias en la resolución de problemas vinculados con bases de datos espaciales y espacio-temporales. En este contexto, proponemos el estudio y el desarrollo de herramientas para la visualización de aplicaciones vinculadas a las bases de datos mencionadas.

2. Línea de investigación

En la línea de investigación se estudia el diseño y desarrollo de índices espacio- temporales, aplicables a diversos escenarios de movimiento (redes, espacios libres de obstáculos, etc.), considerando la geometría como una disciplina marco en la cual se formalizan aspectos propios de los problemas involucrados. En este contexto, se propone el estudio de optimización de estructuras geométricas que están relacionadas con las bases de datos ya mencionadas y al desarrollo de herramientas para la visualización de estructuras geométricas y aplicaciones vinculadas con bases de datos espacio- temporales.

Como objetivos específicos de estudio en la línea de investigación se enumeran los siguientes:

a) Indexación espacio-temporal sobre objetos en movimiento para diversos escenarios. Desarrollo de las estructuras de almacenamiento, los algoritmos de consulta y la evaluación experimental, mostrando el desempeño de los distintos índices en aplicaciones de diferentes magnitudes respecto de la población de objetos en movimiento. Desarrollo de aplicaciones con herramientas de Geometría Computacional y Bases de Datos Espacio-Temporales.

b) Estudio de configuraciones geométricas generales de puntos en el plano considerando medidas de calidad mínimas o máximas, aplicando técnicas metaheurísticas y diversas estrategias algorítmicas.

c) Estudio de problemas de vigilancia considerando variaciones en el tipo de guardia, su posible ubicación y la forma de vigilancia.

d) Diseño y desarrollo de herramientas para la generación, visualización y manipulación de diferentes configuraciones geométricas de conjuntos de puntos en el plano.

3. Resultados Obtenidos /Esperados

Para la resolución de problemas de optimización de Triangulaciones y Pseudo-triangulaciones de Peso Mínimo se aplicaron las técnicas metaheurísticas: Optimización basada en Colonia de Hormigas (Ant Colony Optimization, ACO) y Recocido Simulado (Simulated Annealing, SA), técnicas determinísticas Voraces (Greedy) y Triangulación de Delaunay. Se llevó a cabo el estudio, adecuación y evaluación experimental de las técnicas metaheurísticas mencionadas para la búsqueda de triangulaciones y pseudo-triangulaciones que cumplan la propiedad. Se diseñaron generadores de instancias de problema para ser utilizados en la evaluación experimental.

Los resultados obtenidos para los problemas MWT y MWPT utilizando la técnica

metaheurística ACO fueron publicados en [DGLH11b] [GDLH11] [DGLH12]. Se llevó a cabo una evaluación experimental y análisis de los resultados obtenidos con la técnica Recocido Simulado [DGLH11a] [DGLH11c].

Para el problema Triangulación de Dilación Mínima (Minimum Dilation Triangulation - MDT) todavía no se conoce un algoritmo que lo resuelva en tiempo polinomial y tampoco se ha demostrado que sea NP-duro. Las técnicas aplicadas para este problema fueron: Greedy, Local Search, Iterated Local Search, Simulated Annealing y Random Local Search. Para cada estrategia se propuso un conjunto de operadores adecuados. Debido a la complejidad que implica la puesta a punto de los parámetros de técnicas metaheurísticas, se utilizó Optimización de Parámetros Secuencial (Sequential Parameter Optimization - SPO) para el ajuste de los parámetros requeridos por Simulated Annealing. Se realizó un análisis experimental en el cual se compararon dichos algoritmos con otras técnicas, como por ejemplo, Delaunay. Se crearon las instancias de prueba, ya que para estos problemas no se encontraron disponibles ningún tipo benchmark con el cual comparar nuestros resultados. Las conclusiones fueron afirmadas desarrollando un estudio estadístico aplicando diferentes test estadísticos y métodos de visualización [DLMH14]. Los resultados del tratamiento de estos problemas han dado lugar a tesis de doctorado.

Por otra parte, se desarrolló una herramienta para la generación y visualización de triangulaciones pseudo-triangulaciones y poligonizaciones de conjuntos de puntos en el plano [PDG13].

Además, se implementó una aplicación en el ámbito de la Salud para el seguimiento de focos epidémicos utilizando base de datos espacios-temporales y herramientas de Geometría Computacional [GPDGT14].

El desarrollo de las herramientas antes mencionadas se plasmó en trabajos finales de Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas.

Como trabajo futuro, considerando los problemas sobre configuraciones geométricas, se pretende continuar con el estudio de los problemas de optimización aplicando otras técnicas metaheurísticas, adecuadas para su resolución.

Se comenzará con el estudio de problemas relacionados con otras configuraciones geométricas, tales como cuadrangulaciones, pseudocuadrangulaciones. Se llevarán a cabo, la evaluación de las distintas técnicas para determinar su funcionalidad e impacto en la comunidad científica y el análisis que incluirá el correspondiente tratamiento estadístico, y comparativo con otros algoritmos de tipo aproximado o de tipo exacto.

Por otra parte, con el fin de promover una plataforma de investigación, intercambio y desarrollo de tecnologías de la información y la comunicación (TIC), se integró el Proyecto Campo Conectado. Las organizaciones socias se proponen realizar acciones vinculadas a la promoción de las TIC en el ámbito de la producción agropecuaria con un horizonte inicial a dos años (2017-2018). El objetivo principal es aportar a la cooperación científico tecnológica y a las prácticas sociales, productivas y comerciales de los actores de la producción agropecuaria del semiárido central argentino.

En el marco de Campo Conectado, se propone accionar en la gestación de herramientas y desarrollos tecnológicos aplicados a la gestión de la producción agropecuaria en sistemas reales de producción.

4. Formación de Recursos Humanos

La formación del grupo de trabajo en la Universidad Nacional de San Luis, se consolida con actividades de cooperación mutua e intercambio recíproco de información científica, tecnología y desarrollo de nuevos conocimientos con investigadores locales y de otras universidades.

Entre las actividades más destacadas, se mencionan:

- i) Formación de recursos humanos plasmada en tesis doctorales, tesis de maestría y Licenciados en Ciencias de la Computación.
- ii) Realización y dirección de pasantías de investigación con docentes de otras universidades.
- iii) Actividades de divulgación científica, conferencias y publicaciones en congresos y revistas en el ámbito nacional e internacional.
- iv) Actividades de formación académica, a

través de la realización y dictado de cursos de posgrado y de especialización.

- v) Integración del proyecto Campo Conectado. Como objetivo se propone continuar con las actividades integradoras relacionadas al presente proyecto, proponiendo actividades de formación académica, de formación de recursos humanos locales y de otras universidades nacionales, investigación, desarrollo, y otras actividades académico- científicas vinculantes.

5. Bibliografía

- [BCKO08] de Berg, M., Cheong O., van Kreveld, M., Overmars, M., *Computational Geometry: Algorithms and Applications*. 3rd edition, Springer-Verlag, Heidelberg, 2008.
- [DGLH11a] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Approaches for MWT and MWPT Problems*. XVII Congreso Argentino de Ciencias de la Computación 2011 (CACIC 2011), 2011
- [DGLH11b] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Using ACO metaheuristic for MWT problem*. XXX International Conference of the Chilean Computer Science Society. ISBN 978-0-7695-4689-6. Chile. 2011
- [DGLH11c] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Metaheuristic approaches for MWT and MWPT Problems*. XIV Encuentros de Geometría Computacional. Páginas: 79-82. 2011
- [DGLH12] Dorzán M.G., Gagliardi E.O., Leguizamón M.G., Hernández Peñalver G. *Approximations on Minimum Weight Triangulations and Minimum Weight Pseudo-Triangulations using Ant Colony Optimization Metaheuristic*. Fundamenta Informaticae. ISSN: 0169-2968 (Print), 1875-8681 (Online). Volume 119, number 1, pp 1-27.
- [DLMH14] M. G. Dorzán, M. G. Leguizamón, Efrén Mezura-Montes, G. Hernández Peñalver *Approximated algorithms for the Minimum Dilation Triangulation Problem*. Journal of Heuristics. DOI 10.1007/s10732-014-9237-2. Print ISSN 1381-1231. Online ISSN 1572-9397. Publisher Springer US. 2014.
- [GDLH11] Gagliardi E. O., Dorzán M. G., Leguizamón M. G. y Hernández Peñalver. G.; *Approximations on Minimum Weight Pseudo-Triangulation problem using Ant Colony Optimization*. XXX International Conference of the Chilean Computer Science Society. ISBN 978-0-7695-4689-6. Chile. 2011
- [GPDGT14] Guasch, M.M; Piergallini, M.R ; Dorzán, M.G.; Gagliardi, E.O.; Taranilla, M.T.; *“Una herramienta para el análisis y seguimiento de focos epidémicos”* en Anales del 17º Concurso de Trabajos Estudiantiles en 43 JAIIO. Páginas: 35-45- Universidad de Palermo, Buenos Aires, Argentina, 2014.
- [GL07] Gudmundsson J., Levcopoulos C.; *Minimum weight pseudo-triangulations*. Computational Geometry. Theory and applications. Elsevier Vol. 38- pages 139-153, 2007.
- [MR06] Mulzer W., Rote G. *Minimum weight triangulation is NP-hard*. In Proceedings of the 22nd Annual ACM Symposium on Computational Geometry. 2006.
- [MF04] Michalewicz Z., Fogel D., *How to Solve It: Modern Heuristics*, Springer, 2004.
- [PDG13] Palmero, P.R., Dorzán, M. G., Gagliardi E.O., *Una Herramienta para la Manipulación de Configuraciones Geométricas*, 42º Jornadas Argentinas de Informática e Investigación Operativa - 16º Concurso de trabajos de fin de carrera. Universidad Nacional de Córdoba, Córdoba, Argentina. 2013. ISSN: 1850-2946.
- [TGH11] Taranilla, M. T.; Gagliardi, E. O.; Hernández Peñalver, G. *“Approaching Minimum Area Polygonization”*. Congreso Argentino de Ciencias de la Computación 2011 Páginas: 31-40, 2011.